



# Comparative structural analyses of selected spike protein-RBD mutations in SARS-CoV-2 lineages

Urmi Roy<sup>1</sup>

Received: 25 June 2021 / Accepted: 1 November 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

The severity of COVID-19 has been observed throughout the world as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) globally claimed more than 2 million lives and left a devastating impact worldwide. Recently several virulent mutant strains of this virus, such as the B.1.1.7, B.1.351, and P1 lineages, have emerged with initial predominance in UK, South Africa, and Brazil. Another extremely pathogenic B.1.617 lineage and its sub-lineages, first detected in India, are now affecting some countries at notably stronger spread-rates. The present paper computationally examines the time-based structures of B.1.1.7, B.1.351, and P1 lineages with selected spike protein mutations. The mutations in the more recently found B.1.617 lineage and its sub-lineages are explored, and the implications for multiple point mutations of the spike protein's receptor-binding domain (RBD) are described. The selected S1 mutations within the highly contagious B.1.617.2 sub-lineage, also known as the delta variant, are examined as well.

**Keywords** Mutation · SARS-CoV-2 lineages · Spike protein · Structural biochemistry · Variants, Virus structure

## Introduction

COVID-19, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has generated a devastating global impact, coupled with widespread loss of human lives since March 2020 [1]. In recent months, several variants of this life-threatening virus have been emerged with greater spread-rates, adaptability, and fitness. Among these virulent lineages, B.1.1.7, B.1.351, and P1 (nomenclatures based on the Pango lineage [2]) were initially detected in UK, South Africa, and Brazil, respectively [3]. As summarized in Table 1, these “Variant of Concern” have several notable mutations, some of which are common among the three lineages. The table also includes the highly pathogenic B.1.617 lineage and some of its sub-lineages, that was first found in India more recently, and already included in CDC’s “Variant of Interest” list [4]. The B.1.617.2 sub-lineage (also known as delta variant) found its place in the “Variant of Concern” list on June 15th, 2021. The present work focuses on a set of

comparative structural analyses of these new SARS-CoV-2 variants.

Computational structural biology is rapidly becoming an integral part of applied immunology, as this field continues to aid the understanding of the structural basis of proteins, and thus, plays a key role in the development of preventive drug designs [6–8]. Since the beginning of the COVID-19 pandemic, the availability of experimental results about the structure/function, epidemiological distribution, and mutational fitness of this novel pathogen has been very limited in the commonly available literature. As a result, scientists have heavily relied on simulation-based tools and strategies to investigate this virus. In this regard, computational tools of immunoinformatics can be particularly useful to investigate such evolving infectious pathogens and host-pathogen interactions [9–13]. Our present effort is guided by these considerations.

In previous works of this author, several biologically relevant protein structures as well as mutant models of the angiotensin peptide coordinated to the Zn-bound angiotensin-converting enzyme 2 (ACE2) receptor have been analyzed; more recently, a study of a model structure of the SARS-CoV-2 N501Y variant has also been reported [14–19]. The current investigation of SARS-CoV-2 lineages examines the implications for multiple point mutations on the spike (S)

✉ Urmi Roy  
urmi@clarkson.edu

<sup>1</sup> Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699-5820, USA

**Table 1** Selected mutations in SARS-CoV-2 lineages [3–5]

Name of the lineage/sub-lineage	Spike protein mutations	Experiments with selected mutations on SARS-CoV-2 spike RBD
B.1.1.7 [3]	69/70 deletion, 144 deletion, E484K*, S494P*, N501Y, A570D, D614G, and P681H	E484K, S494P, and N501Y
B.1.351 [3]	K417N, E484K, N501Y, and D614G	K417N, E484K, and N501Y
P1 [3]	K417T**, E484K, N501Y, and D614G	K417T, E484K, and N501Y
B.1.617 [4]	L452R, E484Q, and D614G	L452R and E484Q
B.1.617.1 [4]	T95I*, G142D, E154K, L452R, E484Q, D614G, P681R, and Q1071H	
B.1.617.3 [4]	T19R, G142D, L452R, E484Q, D614G, P681R, and D950N	
Variant with possible mutations [5]	L452R and Y453F	L452R and Y453F
Name of the lineage/sub-lineage	Spike protein mutations	Experiment with selected mutations on SARS-CoV-2 spike S1
B.1.617.2 [3]	T19R, (G142D), 156/157 deletion, R158G, L452R, T478K, D614G, and P681R	L452R, T478K, and D614G

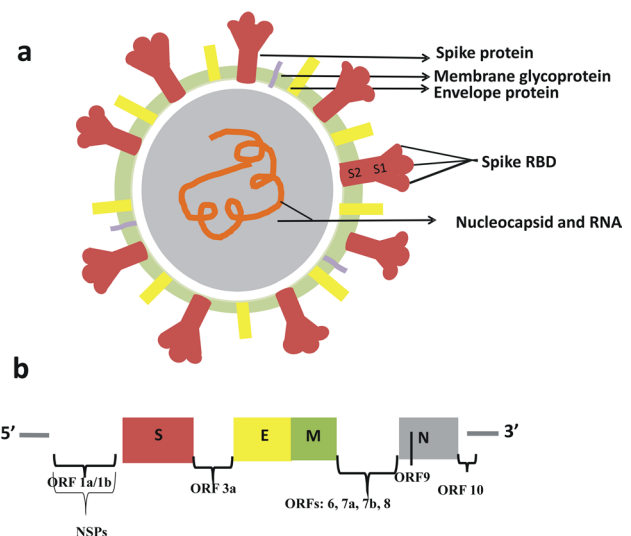
\*Found in some sequences

\*\*Initially observed as K417N/T, later this mutation was identified as K417T in the P1 lineage

RBD. In particular, this study will measure the structural and conformational variations of these mutant variants as functions of time and demonstrate how structural change corresponds to their functions. The B.1.617.2 sub-lineage, namely, the delta variant, is the prevalent form of this virus; the structural impacts of selected S1 mutations within the delta variant are also discussed here. In remainder of this report, the B.1.617.2 sub-lineage will be simply referred to as delta variant.

The SARS-CoV-2 genome contains non-structural (NSPs) as well as structural proteins. There are 16 NSPs in SARS-CoV-2 genome, NSP1–NSP16. The structural portions consist of spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins. The S glycoprotein is the main interacting site for the host entry and plays an important role in host defense and antibody neutralization. The accessory proteins contain several open reading frames (ORFs) including ORF1a/1b, ORF3a/3b, ORF6, ORF7a/7b, and ORFs8–10. The SARS-CoV-2 genome contains a single strand sense RNA, containing some similarities with the previously identified beta-coronavirus family members of the SARS-CoV and the Middle East respiratory syndrome (MERS-CoV) viruses.

Illustrations of SARS-CoV-2 and its genome structure are schematically presented in Fig. 1a and b. Though SARS-CoV-2 has some structural similarities with SARS and MERS, as observed in their sequence-based similarities, the fast transmissibility and adaptability of the highly pathogenic SARS-CoV-2 is rather unique. The long S protein of SARS-CoV-2 contains ~1273 amino acids (AAs) with a small N-terminal signaling peptide (~1–13 AAs)



**Fig. 1** **a** Schematic diagram of the SARS-CoV-2 structure where the spike proteins are displayed in light brown. **b** The genome structure of the SARS-CoV-2

at the beginning of the sequence. The protein continues with the S1 (~14–685 AAs) and S2 (~686–1273 AAs) subunits. The S1 subunit is essential for receptor binding and is comprised of two domains, N-terminal domain (NTD; ~14–305 AAs) and receptor-binding domain (RBD; ~319–541 AAs) [20]. The modeling/mutation schemes of the protein used in this study are based on the PDB structure (6M0J:E) and are centered on the S1 RBD. The truncated form of 6Z97.PDB structure is used for the delta variant with selected S1 mutations.

## Materials and methods

The S protein RBD was used for the first set of simulations reported here. The wt type 6M0J: E as SARS-CoV-2 S1 RBD [21] and some of the known mutations found in the B.1.1.7, B.1.351, and P1 lineages were selected. The B.1.617, lineage, as well as the sub-lineages B.1.617.1/B.1.617.3 with selected common mutations in their S1-RBD were analyzed. A possible mutant with two single-point mutations in its S protein's RBD was also examined [5]. The second set of simulations considered in this work involved the S1. The truncated version of the 6Z97 (319 to 620 AAs) was considered as wt S1 [22], and for the delta variant, the selected mutations within the S1 subunit were chosen.

All these simulations used the Nanoscale Molecular Dynamics (NAMD), quickMD, and Visual Molecular Dynamics (VMD) software programs [23–25]. The proteins' 3D models were set up by using Biovia's Discovery Studio Visualizer [26].

## Results and discussion

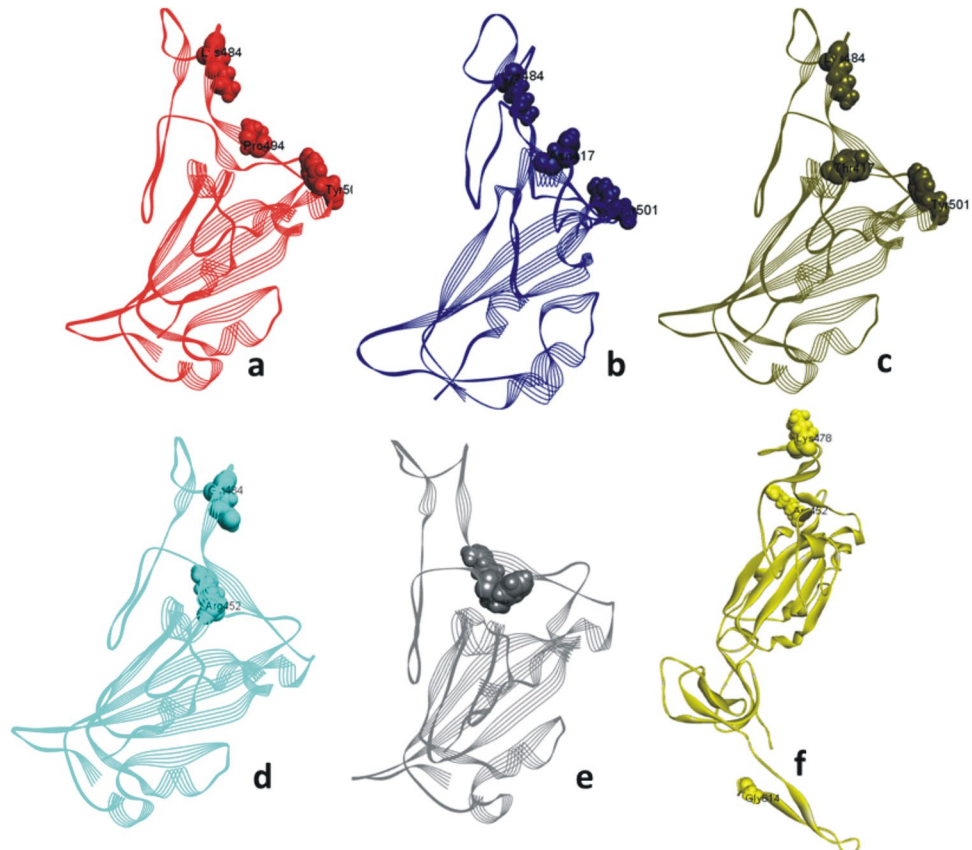
In total, eight simulations were executed for the mutations listed in Table 1, one wt S1 RBD, and five S1-RBD variants with selected mutations. Two supplementary simulations

were performed, with wt S1 and the delta variant with selected S1 mutations. Starting from the native pdb structures, 6M0J:E and 6Z97:A, the mutant variants were generated using the mutator gui of Visual Molecular Dynamics (VMD) [23]. After completing the initial protocols (minimization/annealing and equilibration processes), MD simulation was continued for 30 ns. The integration time was 2 fs for all procedures. All protocols used the generalized Born solvent-accessible surface area implicit solvation model [27]. For annealing and equilibration, the backbones were restrained, but no atoms were restrained during the final simulation process. Using Langevin dynamics, the temperature was maintained at 300 K during final simulation process. The details of the simulation protocols are described elsewhere [19].

Fig. 1a shows a generalized schematic of the wt SARS-CoV-2 structure, where the S, E, M, N, and the viral RNA, as well as the two subunits S1 and S2 of the S protein are identified. The mutations considered in this report are found in the RBD and the S1 subunit of the S protein. Fig. 1b shows a typical genomic display of the SARS-CoV-2, where different nonstructural and structural parts are presented, along with the open reading frames (ORFs).

As shown in Fig. 2a–e, the structural illustrations of the wt and mutant variants of the SARS-CoV-2 RBD correspond to the 6M0J:E subunit. This is a relatively small (333–526)

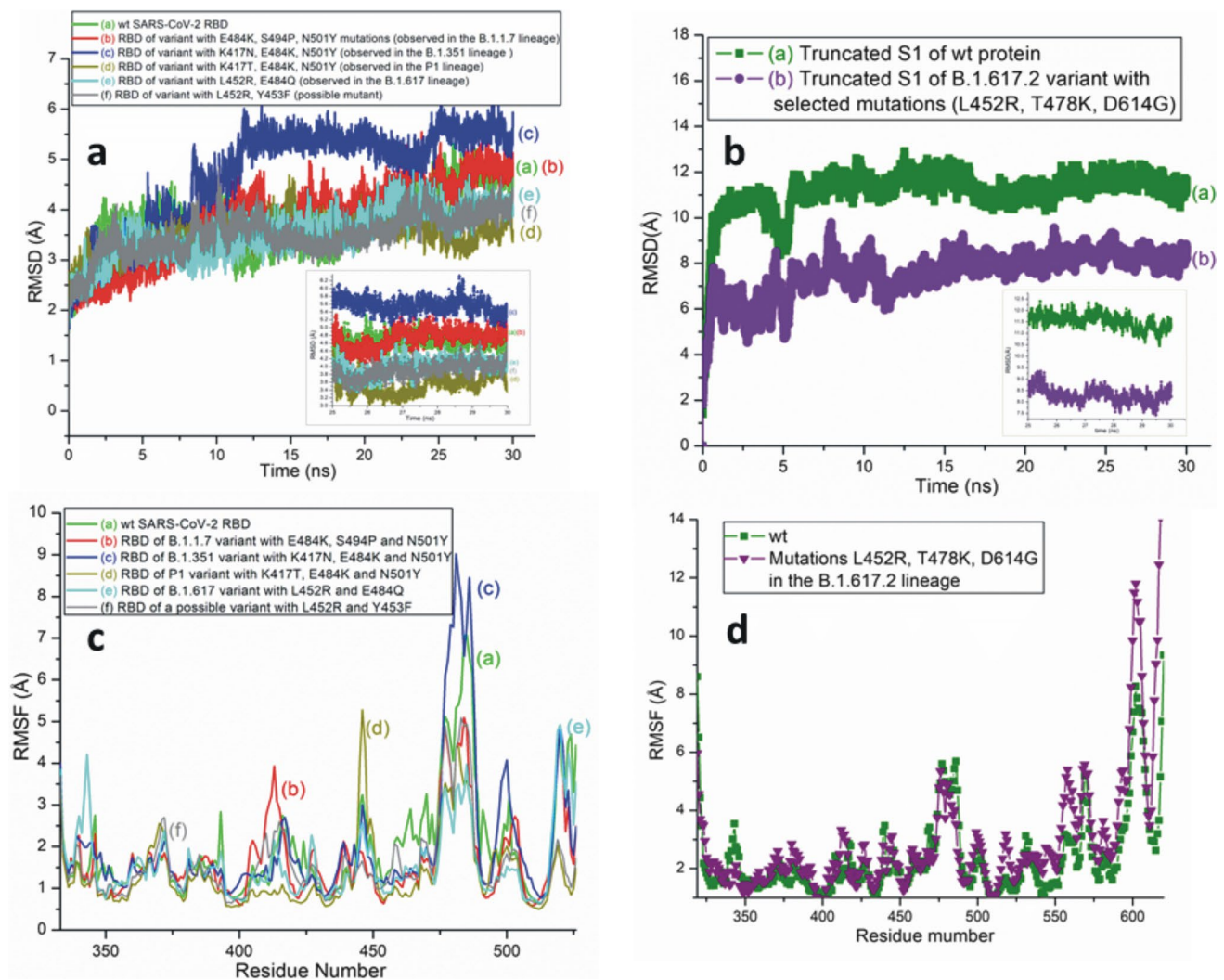
**Fig. 2** Secondary structures of SARS-CoV-2: S1-RBD/S1 variants with selected mutations. **a–e** The mutant RBD is based on 6M0J.PDB. **f** The mutant S1 is based on the truncated 6Z97. PDB. **a** The selected mutant residues E484K, S494P, and N501Y of B.1.1.7 lineage. **b** The selected mutant residues K417N, E484K, and N501Y of B.1.351 lineage. **c** The selected mutant residues K417T, E484K, and N501Y of P1 lineage. **d** The mutant residues L452R and E484Q of B.1.617 lineage. **e** The mutations L452R and Y453F within a possible mutant structure. **f** The selected signature mutant residues L452R, T478K, and D614G within the B.1.617.2 sub-lineage, the delta variant



subunit, and from Fig. 2a–c it is evident that some of these mutations occur on the surface of the protein. Selected delta variant mutations on S1 are displayed in Fig. 2f. These three mutated S1 residues are also exposed residues. It should be noted in this context that during the second wave of COVID-19 observed in India, the newly identified sub-lineages B.1.617.1 and B.1.617.2 were the most prevalent forms [28]. The B.1.617.2 sub-lineage, the delta variant, has a different mutation, T478K instead of E484Q that is observed in sub-lineages B.1.617.1 and B.1.617.3. Three signature S1 mutations L452R, T478K, and D614G possessed by the delta variant have been included in this work. Due to its rather recent emergence, some of the mutation-containing regions of the delta variant, and in particular the structure close to

the furin cleavage site connecting the S1/S2 subunits, are still missing in the commonly available published structures of this species.

Fig. 3a shows comparative RMSD plots of the SARS-CoV-2 variants with selected mutations based on the RBD of 6MOJ.PDB. All these variants reached convergence, during the last 5 ns of the simulation. Nevertheless, as observed in the inset figure, the selected mutations with B.1.351 lineage (c) show measurably higher RMSD values than those of P1 (d), B.1.617 (e), and the species of combined mutations 452 and 453 (f) during the last phase of the simulation. These last three variants (RMSD values; d: 3.53 Å; e: 3.99 Å; and f: 3.91 Å) are more stable than their wt species (a: 4.66 Å). The selected mutations within these variants are displayed in



**Fig. 3** **a** RMSD plots of wt and different SARS-CoV-2 variants with selected S1-RBD mutations. The two mutations identified in B.1.617 lineage (e) are also common in B.1.617.1 and B.1.617.3 sub-lineages. Inset showing the RMSDs for the last 5 ns. **b** The RMSD graphs of wt S1 and the B.1.617.2 variant with selected S1 mutations. Inset dis-

play the RMSDs for the last 5 ns. **c** The RMSF plots for the wt and different SARS-CoV-2 variants with selected S1-RBD mutations. **d** The RMSF plots for the wt S1 and B.1.617.2 variant with selected S1 mutations



Supplementary Information (SI) Fig. S1, where once again, the selected mutations of P1, B.1.617, and the third variant are seen to exhibit the lowest RMSD values, indicating a fairly stable nature shared by these mutations. The foregoing plots demonstrate that the selected mutations within P1, B.1.617 (also B.1.617.1 and B.1.617.3) lineages, and the possible variant with 452 and 453 mutations are more stable than the mutant variants of B.1.351 and B.1.1.7. It is unknown, however, if the protein's stability is dictated by the number of mutations within the variant.

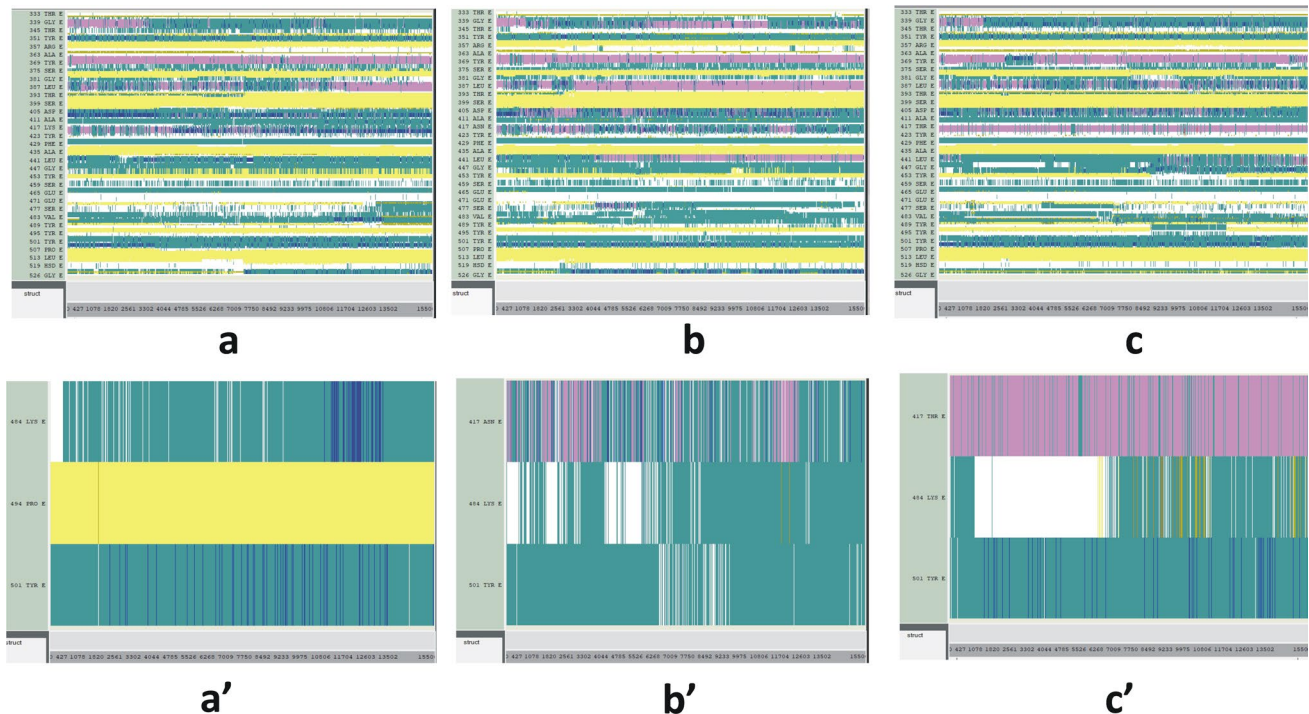
Fig. 3b presents RMSD plots of the wt S1 and the S1 of delta variant with selected mutations. The S1 is based on the truncated form of 6Z97.PDB. From Fig. 3b inset, it is evident that the RMSD plot is quite stable and that the averaged value of the all-atom RMSD during the last 5 ns within the delta variant (8.3 Å) is much lower than its wt version (11.54 Å). The overall higher value of S1 RMSD than the RBD RMSD is linked to the presence of many turns and loops within the S1 structures.

As in the case of the RMSD data, the RMSF plots in Fig. 3c also show that the B.1.617 lineage (e) and the L452R-Y453F mutations (f) are characterized by minimal fluctuations (lowest in their comparison group). At the same time, the selected mutations in B.1.351 lineage (c) once again show higher fluctuations in Fig. 3c. While the

RMSD and RMSF graphs have already been plotted for the wt RBD structure, for a direct comparison, they have been included in Fig. 3a and c [19]. In Fig. 3d, the RMSF values in L452R and T478K are somewhat lower than their wt residues, although the RMSF for the D614G is rather high. This shows that the two mutant residues, L452R and T478K, are steadier (more stable) than D614G. SI Fig S2 shows the hydrogen bond numbers during the simulation time; for none of the cases considered, these numbers exhibit any significant variations.

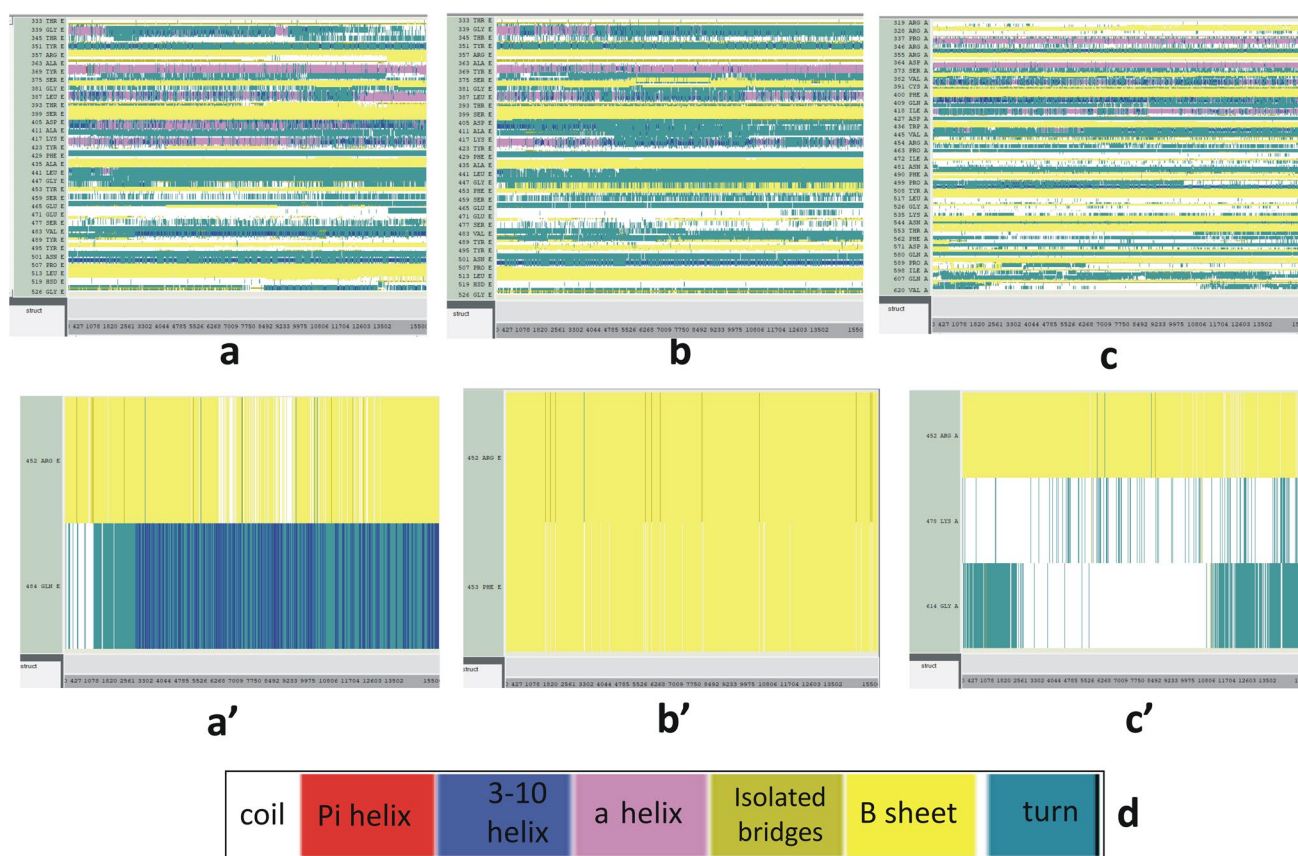
Figs. 4 and 5 represent time-based secondary structure changes of the variant proteins as well as those of the selected mutant residues within these mutant strains. Secondary structures, in particular, the  $\alpha$  helices and  $\beta$  sheets play a crucial role in determining proteins stability. From Fig. 4a-a', it is clear that the mutation, S494P in B.1.1.7, is rather stable. Near the end of the simulation time, the E484K and N501Y transformations show slight fluctuations from coils to  $3_{10}$  helices though they disappeared at the end of the simulation. However, the actual manifestation of these latter effects may change with the progression of time.

Within the B.1.351 variant in Fig. 4b-b', the K417N mutation shows some variations from  $\alpha$  helices to  $3_{10}$  helices, turns and coils. This makes the K417 unstable during the last phase with the extinction of the  $\alpha$  helices. The



**Fig. 4** Time-based secondary structure changes of SARS-CoV-2 mutant S1-RBDs. The time-based structure changes of **a** the variant of SARS-CoV-2 RBD with mutations E484K, S494P, and N501Y. **a'** E484K, S494P, and N501Y residues within the B.1.1.7 lineage. **b** The variant of SARS-CoV-2 RBD with mutations K417N, E484K,

and N501Y. **b'** The K417N, E484K, and N501Y residues within the B.1.351 lineage. **c** The variant of SARS-CoV-2 RBD with mutations K417T, E484K, and N501Y. **c'** K417T, E484K, and N501Y residues within the P1 lineage



**Fig. 5** Secondary structure changes of **a** the variant of SARS-CoV-2 RBD with mutations L452R and E484Q. **a'** L452R and E484Q residues within the B.1.617/B.1.617.1/B.1.617.3 lineages. **b** The variant of SARS-CoV-2 RBD with mutations L452R and Y453F. **b'** L452R

and Y453F residues within a possible mutant variant. **c** The variant of SARS-CoV-2 with selected S1 mutations L452R, T478K, and D614G. **c'** L452R, T478K, and D614G residues within the delta variant. **d** The color code of secondary structure analyses

E484K mutation in this variant may show higher level of stability as some coils and turns are converted to few  $\beta$  sheets near the last phase of the simulation—even though they eventually disappear at the end. Coils to turns are still observed here at the end of simulation. There are some changes from turns to coils in the 501 residue, but it is not clear from the data if the latter correspond to the so-called coiled coils.

Among the mutations we study here in P1, the K417T and N501Y do not show any significant structure variations throughout the simulation time. The E484K mutation exhibits a higher level of stability as some coils and turns in this species are converted to  $\beta$  sheets during the last phase of the simulation (Fig. 4c-c').

Within the L452R residue of B.1.617 lineage and the B.1.617.1/B.1.617.3 sub-lineages, the intermittent 3-10 helices and the isolated bridges are completely transferred into  $\beta$  sheets making this mutant residue more stable. In the case of E484Q, the random coils and turns are converted to  $3_{10}$  helices that are more stable, which, consequently make the secondary structure more rigid and solid (Fig. 5a-a'). The L452R and Y453F mutations in Fig. 5b-b' are stable

overall and do not indicate any significant secondary structure changes during simulation.

Fig. 5c-c' describes the secondary structure changes of the delta variant with selected S1 mutations L452R, T478K, and D614G. Here L452R is stable as it mostly consists of  $\beta$  sheets. Mutation T478K is also stable as some  $\beta$  sheets are observed during the last phase of the simulation. The T478K is a unique mutation that is only observed within the delta variant. The other sub-lineages of B.1.617 do not exhibit this specific mutation and, to our knowledge, this mutation is not seen in any other lineages identified so far. During the last part of the simulation, mostly turns are observed in D614G and there are no major changes from its initial phase. The default color code of proteins' secondary structure analyses is displayed in Fig. 5d.

The secondary structure change of a protein is a key factor necessary to understand the latter's conformational changes, and such changes might act to affect their functionality. From Figs. 4 and 5, it is evident that the selected mutant residues within the B.1.351 strain are fairly unstable, and that the residues within the B.1.617/B.1.617.1-B.1.617.3

variants are mostly stable. Among these mutations, L452R, Y453F, E484Q, and S494P are quite stable, and T478K and E484K is particularly stable within the delta and P1 strains, respectively.

While numerous studies have already been reported on the structure and functions of SARS-CoV-2, the structures of the more recently found lineages have not been thoroughly investigated yet. The S1-RBD mutations of the B.1.617 lineage and its associated sub-lineages are of particular interest in this context. The B.1.617 strain was initially labeled as “double mutant” due to the presence of two mutations (E484Q, L452R) from two different lineages; L452R is found in B.1.427/B.1.429 while E484K exists in both B.1.351 and P1. However, in B.1.617 the E484K mutation has been changed to E484Q. Another possible mutant variant with mutations L452R and Y453F is the combination of two lineages, B.1.427/B.1.429 and B.1.298. The stable mutation in residue 452 may form a stronger complex with ACE2. The Y453F in the Fig. 5b’ also exhibits stronger stability with time.

The wt E484 residue has been recognized as a “repulsive residue” between the RBD-ACE2 complex [29]. Since the mutations in the E484K/Q residue are particularly stable in the P1 and B.1.617 strains (Fig. 4c and 5a), the mutations in this residue may form a stronger bond with the receptor. S494P is also very strong as it resides within the  $\beta$  helix; within the B.1.1.7 lineage this appears to be the most stable residue. There are numerous examples in the literature that N501Y mutation forms a stable connection with the receptor [19, 30]. Likewise the mutations L452R and T478K are also strongly bound to the ACE2 receptor [28]. Nonetheless, mutation D614G plays a major role in preventing premature binding of S1 to the ACE2 receptor as well as the dissociation of S1 subunit within the trimeric framework; thus, the D614G promotes greater transmissibility [31, 32].

It could be possible that some of these aforesaid mutations contribute to stability changes as a result of changes in their charges and/or hydrophobicity [33]. Most of the mutations examined in this study are surface exposed. Within the delta variant mutation L452R, the hydrophobic Leu may be unstable in an exposed environment whereas the positively charged hydrophilic Arg tends to be more stable and may be more interactive with the receptor protein. For T478K, the Thr is a polar non-charged residue whereas the Lys is hydrophilic and positively charged. Therefore, Lys may be more interactive than Thr. In D614G, the negatively charged hydrophilic Asp is mutated to hydrophobic Gly. As residue 614 is also surface exposed, the hydrophobic Gly is more unstable than the hydrophilic Asp. Thus, mutation D614G is more unstable than the other two mutations L452R and T478K.

While a stable RBD and S1 may be important for accurate protein enthalpy, stability variations (with respect to wt) are

largely dictated by mutations within the interaction surface. Furthermore, protein-protein binding may require a certain level of flexibility. Accordingly, any changes of stability may alter the process and may explain the observed differences of virulence between the lineages [34]. The mutational analysis available through COVID3D would be another possible approach to mapping the structural consequences of mutations within different SARS-CoV-2 lineages [35].

Existing literature suggests that some of the stable mutations studied here may act as partial neutralizing escape mutations with lower antibody binding affinity. As for example, L452R shows reduced antibody affinity and E484K is generally recognized as mAb escape mutation. Moreover, in vitro experiment suggested that mutation T478K may escape from the immune recognition process [33, 36–38].

## Conclusions

According to the results presented here, the mutant RBD variant of B.1.617 (as well as some of its sub-lineages), P1, and the potential variant with two possible mutants are stable forms. The delta variant with selected S1 mutation is also fairly stable (Fig. 3a–b). The steadier and lower average RMSD values of the above-mentioned variants compared to those of their wt species are particularly indicative of the former’s stable nature. Among the mutations we have studied in this work, L452R, Y453F, T478K, E484Q, and S494P are relatively stable. N501Y does not show significant variations during the simulation timescale. The E484K within the P1 strain is also fairly stable. Since these newly found lineages are more spreadable than their predecessor species, some of the stable mutations may escape from the antibody neutralization and cellular immunity. In fact, some of the variants with viral mutations K417N/T, L452R, Y453F, T478K, and E484K are recognized as immune escape mutants [36–38].

The stable mutations found here to occur within the highly infective species may help to further understand for the associated antibody cross-reactivity and may also facilitate the task of designing effective inhibitors. A correlation between protein stability and the virulence between different lineages has also been noted in this study. The computational investigation presented here may contribute to the ongoing efforts to expand the knowledgebase currently available about COVID-19. Specifically, the enhanced stabilities of some of the mutant residues, as those found here for the newer variants, may have implications in the context of future vaccine developments to combat other impending strains and pathogenic variants of SARS-CoV-2.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12026-021-09250-z>.



**Acknowledgments** The author acknowledges utilization of the following simulation and visualization software packages: (1) NAMD and (2) VMD: NAMD and VMD, developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois, Urbana-Champaign. (3) Discovery Studio Visualizer: Dassault Systèmes BIOVIA, Discovery Studio Modeling Environment, San Diego, CA: Dassault Systèmes (2015).

## Declarations

**Conflict of interest** The author declares no competing interests.

## References

- WHO. Weekly epidemiological update. 2021. <https://www.who.int/publications/m/item/weekly-epidemiological-update---23-february-2021>. Accessed 25 Feb 2021.
- Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol*. 2020;5(11):1403–7. <https://doi.org/10.1038/s41564-020-0770-5>.
- Centers for Disease Control and Prevention. SARS-CoV-2 variant classifications and definitions. <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#Concern>. Accessed 16 Mar 2021.
- Centers for Disease Control and Prevention. SARS-CoV-2 variant classifications and definitions. <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#Interest>. Accessed 5 May 2021.
- L452R and Y453F SARS-CoV-2 mutations increase transmission and evade immunity. <https://www.news-medical.net/news/20210411/L452R-and-Y453F-SARS-CoV-2-mutations-increase-transmission-and-evade-immunity.aspx>. Accessed April 19, 2021.
- Haseeb M, Anwar MA, Choi S. Molecular interactions between innate and adaptive immune cells in chronic lymphocytic leukemia and their therapeutic implications. *Front Immunol*. 2018;9:2720. <https://doi.org/10.3389/fimmu.2018.02720>.
- Head JF, Swamy N, Ray R. Crystal structure of the complex between actin and human vitamin D-binding protein at 2.5 Å resolution. *Biochemistry*. 2002;41(29):9015–20. <https://doi.org/10.1021/bi026054y>.
- Webb RN, Cruse JM, Lewis RE. Decreased TLR4 gene expression in leukemic leukocyte populations. *Exp Mol Pathol*. 2009;87(2):117–26. <https://doi.org/10.1016/j.yexmp.2009.07.007>.
- Sheinerman FB, Giraud E, Laoui A. High affinity targets of protein kinase inhibitors have similar residues at the positions energetically important for binding. *J Mol Biol*. 2005;352(5):1134–56. <https://doi.org/10.1016/j.jmb.2005.07.074>.
- Gray NW, Zhorov BS, Moczydlowski EG. Interaction of local anesthetics with the K (+) channel pore domain: KcsA as a model for drug-dependent tetramer stability. *Channels (Austin)*. 2013;7(3):182–93. <https://doi.org/10.4161/chan.24455>.
- Salopek-Sondi B, Swartz D, Adams PS, Luck LA. Exploring the role of amino acid-18 of the leucine binding proteins of *E. coli*. *J Biomol Struct Dyn*. 2002;20(3):381–7. <https://doi.org/10.1080/07391102.2002.10506856>.
- Bello M, Campos-Rodriguez R, Rojas-Hernandez S, Contis-Montes de Oca A, Correa-Basurto J. Predicting peptide vaccine candidates against H1N1 influenza virus through theoretical approaches. *Immunol Res*. 2015;62(1):3–15. <https://doi.org/10.1007/s12026-015-8629-1>.
- Kanduc D, Shoenfeld Y. Molecular mimicry between SARS-CoV-2 spike glycoprotein and mammalian proteomes: implications for the vaccine. *Immunol Res*. 2020;68(5):310–3. <https://doi.org/10.1007/s12026-020-09152-6>.
- Roy U. Structural characterizations of the fas receptor and the fas-associated protein with death domain interactions. *Protein J*. 2016;35(1):51–60. <https://doi.org/10.1007/s10930-015-9646-6>.
- Roy U. Structural modeling of tumor necrosis factor: a protein of immunological importance. *Biotechnol Appl Biochem*. 2017;64(4):454–63. <https://doi.org/10.1002/bab.1523>.
- Roy U. 3D modeling of tumor necrosis factor receptor and tumor necrosis factor-bound receptor systems. *Mol Inform*. 2019;38(5):e1800011. <https://doi.org/10.1002/minf.201800011>.
- Roy U. Structural and molecular analyses of functional epitopes and escape mutants in Japanese encephalitis virus envelope protein domain III. *Immunol Res*. 2020;68(2):81–9. <https://doi.org/10.1007/s12026-020-09130-y>.
- Fatouros PR, Roy U, Sur S. Modeling substrate coordination to Zn-bound angiotensin converting enzyme 2. *bioRxiv*. 2021. <https://doi.org/10.1101/2021.03.27.437352>.
- Roy U. Role of N501Y mutation in SARS-CoV-2 spike protein-structure. Preprints, 2021060238 2021. <https://doi.org/10.20944/preprints202106.0238.v1>
- Huang Y, Yang C, Xu X-f, Xu W, Liu S-w. Structural and functional properties of SARS-CoV-2 spike protein: potential anti-virus drug development for COVID-19. *Acta Pharmacol Sin*. 2020;41(9):1141–9. <https://doi.org/10.1038/s41401-020-0485-4>.
- Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*. 2020;581:215–20. <https://doi.org/10.1038/s41586-020-2180-5>.
- Huo J, Zhao Y, Ren J, Zhou D, Duyvesteyn HME, Ginn HM, et al. Neutralization of SARS-CoV-2 by destruction of the prefusion spike. *Cell Host Microbe*. 2020;28(3):445–54.e6. <https://doi.org/10.1016/j.chom.2020.06.010>.
- Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14(1):33–8. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
- Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005;26(16):1781–802. <https://doi.org/10.1002/jcc.20289>.
- Ribeiro JV, Bernardi RC, Rudack T, Stone JE, Phillips JC, Freddolino PL, et al. QwikMD- Integrative molecular dynamics toolkit for novices and experts. *Sci Rep*. 2016;6:26536. <https://doi.org/10.1038/srep26536>.
- Dassault Systèmes BIOVIA. Discovery studio modeling environment. San Diego: Dassault Systèmes; 2015.
- Tanner DE, Phillips JC, Schulten K. GPU/CPU algorithm for generalized born/solvent-accessible surface area implicit solvent calculations. *J Chem Theory Comput*. 2012;8(7):2521–30. <https://doi.org/10.1021/ct3003089>.
- Cherian S, Potdar V, Jadhav S, Yadav P, Gupta N, Das M, et al. SARS-CoV-2 spike mutations, L452R, T478K, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. *Microorganisms*. 2021;9(7):1542. <https://doi.org/10.3390/microorganisms9071542>.
- Yang Y, Zhang Y, Qu Y, Zhang C, Liu XW, Zhao M, et al. Key residues of the receptor binding domain in the spike protein of SARS-CoV-2 mediating the interactions with ACE2: a molecular dynamics study. *Nanoscale*. 2021;13(20):9364–70. <https://doi.org/10.1039/d1nr01672e>.
- Barton MI, MacGowan SA, Kutuzov MA, Dushek O, Barton GJ, van der Merwe PA. Effects of common mutations in the SARS-CoV-2 spike RBD and its ligand the human ACE2 receptor on



- binding affinity and kinetics. *Elife*. 2021;10. <https://doi.org/10.7554/eLife.70658>.
31. Choe H, Farzan M. How SARS-CoV-2 first adapted in humans. *Science*. 2021;372(6541):466–7. <https://doi.org/10.1126/science.abi4711>.
  32. Zhang J, Cai Y, Xiao T, Lu J, Peng H, Sterling SM, et al. Structural impact on SARS-CoV-2 spike protein by D614G substitution. *Science*. 2021;372(6541):525–30. <https://doi.org/10.1126/science.abf2303>.
  33. Di Giacomo S, Mercatelli D, Rakhimov A, Giorgi FM. Preliminary report on severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike mutation T478K. *J Med Virol*. 2021;93(9):5638–43. <https://doi.org/10.1002/jmv.27062>.
  34. MacGowan SA, Barton MI, Kutuzov M, Dushek O, van der Merwe PA, Barton GJ. Missense variants in human ACE2 modify binding to SARS-CoV-2 spike. *bioRxiv*. 2021:2021.05.21.445118. <https://doi.org/10.1101/2021.05.21.445118>.
  35. Portelli S, Olshansky M, Rodrigues CHM, D'Souza EN, Myung Y, Silk M, et al. Exploring the structural distribution of genetic variation in SARS-CoV-2 with the COVID-3D online resource. *Nat Genet*. 2020;52(10):999–1001. <https://doi.org/10.1038/s41588-020-0693-3>.
  36. Greaney AJ, Starr TN, Barnes CO, Weisblum Y, Schmidt F, Caskey M, et al. Mapping mutations to the SARS-CoV-2 RBD that escape binding by different classes of antibodies. *Nat Commun*. 2021;12(1):4196. <https://doi.org/10.1038/s41467-021-24435-8>.
  37. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol*. 2021;19(7):409–24. <https://doi.org/10.1038/s41579-021-00573-0>.
  38. Starr TN, Greaney AJ, Dingens AS, Bloom JD. Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016. *Cell Rep Med*. 2021;2(4):100255. <https://doi.org/10.1016/j.xcrm.2021.100255>.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.